



آزمون بسندگی سنجش زبان پارسی (آسپا): طراحی و بررسی ساختار، پایایی، و روایی^۱

معصومه استاجی^۲، اعظم رستمی^۳

تاریخ دریافت: ۱۴۰۰/۰۴/۰۶

تاریخ پذیرش: ۱۴۰۰/۱۱/۰۵

چکیده

گسترش آموزش زبان فارسی در مؤسسه‌ها و مراکز دانشگاهی مختلف در ایران و تقاضای فارسی‌آموزان خارج از ایران برای شرکت در این دوره‌ها به دلایل گوناگونی، دست‌اندرکاران امر را بر آن داشت تا شیوه واحدی را برای سنجش میزان دانش زبانی فارسی‌آموزان ارائه کنند. این امر سبب طراحی آزمون‌های سامفا و آمفا شد. در این راستا، مرکز زبان فارسی دانشگاه الزهراء بنا به اعلام نیاز یکی از مراکز آموزش زبان فارسی، اقدام به طراحی یک آزمون بسندگی با نام آسپا نمود. مقاله حاضر، به بررسی زوایای مختلف این آزمون و میزان پایایی و روایی آن می‌پردازد. تعداد فارسی‌آموزان شرکت‌کننده در این مطالعه ۴۱ نفر از دانشجویان دوره‌های آموزش زبان فارسی در شهر مشهد بودند. یافته‌های این آزمون نشان می‌دهد که بخش واژگان و دستور از سطح آسانی بیشتر و نوشتن از سطح سختی بالاتری در جامعه آماری مورد بررسی برخوردار است. لازم به گفتن است که برای سنجش پایایی این آزمون از روش آلفای کرونباخ و به منظور سنجش میزان روایی محتوایی آن از دیدگاه‌های ۲۰ کارشناس آموزش زبان فارسی به غیرفارسی‌زبانان بهره گرفته شد. نتایج این بررسی نمایانگر آن است که این آزمون دارای پایایی و روایی محتوایی مناسبی است. بررسی شاخص روایی محتوایی این

^۱ شناسه دیجیتال (DOI): 10.22051/JLR.2022.36747.2055

^۲ دانشیار زبان‌شناسی کاربردی، گروه زبان و ادبیات انگلیسی، هیأت علمی دانشگاه علامه طباطبائی، تهران، ایران (نویسنده مسئول)؛ mestaji74@gmail.com، estaji@atu.ac.ir

^۳ دانشجوی دکتری و مدرس آموزش زبان فارسی به غیرفارسی‌زبانان، دانشگاه علامه طباطبائی، تهران، ایران؛ rostami.azfa@gmail.com

آزمون نشان می‌دهد که مهارت‌های دستور و نوشتن با کسب ۹۰٪ دیدگاه مساعد کارشناسان، بالاترین میزان و در پرسش‌های مربوط به مهارت خواندن کمترین روایی محتوایی را دارد. بنابراین، می‌توان نتیجه گرفت که آسیا به عنوان آزمون سنجش سطح بسندگی زبان فارسی از پایایی و روایی مطلوبی برخوردار است و می‌توان از آن برای سنجش مهارت‌های زبانی فارسی‌آموزان استفاده کرد.

واژه‌های کلیدی: آموزش زبان فارسی به خارجی‌ان، آزمون سنجش پارسی (آسیا)، پایایی آزمون، روایی محتوایی آزمون

۱. مقدمه

آزمون، به هر نوع ابزار یا شیوه‌ای گفته می‌شود که برای سنجش عملکرد، توانایی یا دانش شخص در یک زمینه مشخص به کار می‌رود. آزمون در درجه اول یک روش است، این روش در واقع نوعی ابزار است که اجزاء تشکیل‌دهنده آن مجموعه‌ای از تکنیک‌ها، روش‌های انجام کار و مواد آزمودنی هستند. روش مورد اشاره عموماً مستلزم رفتار یا فعالیتی از سوی آزمودنی یا آزمون‌گیرنده یا هردوی آن‌ها است (Brown, 1980). در این راستا، آزمون بسندگی زبان، آزمون معیاری برای سنجش دانش زبان مقصد است. به ویژه برای اشخاصی که می‌خواهند در کشورهای زبان مقصد درس بخوانند، کار کنند، و یا به آن کشورها مهاجرت نمایند. این آزمون چهار مهارت زبانی نوشتن، خواندن، شنیدن، صحبت کردن و مولفه‌های مرتبط دیگری مانند واژگان و دستور زبان را مورد ارزیابی قرار می‌دهد.

در آزمون بسندگی مهم نیست فرد کجا، چه زمانی، و با چه شیوه و محتوای آموزشی زبان مقصد را آموخته است، بلکه مهارت‌های زبانی وی در زمان برگزاری آزمون ارزیابی و سنجیده می‌شود (Abbasi, 2014). آزمون بسندگی به دو بخش دسته‌بندی می‌شود: آزمون بسندگی برای اهداف خاص^۱ که می‌کوشد سطح و نوع زبان مورد نیاز برای به کار بردن در حوزه‌ای ویژه مانند علوم انسانی، تجربی و موارد مشابه را مشخص کند؛ و آزمون بسندگی عمومی^۲ که مرتبط به موضوع خاصی نبوده و کلی است. طراحی و اجرای این آزمون‌ها، به سبب پیچیدگی مفهوم بسندگی است که به آسانی قابل تعریف نیست و در مقایسه با دیگر آزمون‌ها دشوارتر است. میزان توانش فارسی‌آموزان در زبان فارسی مستقل از کتاب یا دوره درسی گذرانده شده، بسندگی تعریف می‌شود (Vakilifard, 2015). آزمون بسندگی فارسی، داوطلبان بسیاری در ایران و

¹ proficiency test for specific purposes

² general proficiency test

جهان دارد. ولی متأسفانه این آزمون آن‌چنان که باید و شاید هنوز به صورت استاندارد برگزار نمی‌شود (Vakilifard, 2015).

امروزه، با وجود مراکز بسیاری که در داخل و خارج از ایران به آموزش زبان فارسی مشغول هستند، آزمون‌های معیار برای سنجش توانایی زبانی فارسی‌آموزان بسیار اندک است. درحالی‌که آزمون معتبری مانند تافل از سال ۱۹۶۴ که برای نخستین بار برگزار شد تاکنون میلیون‌ها شرکت‌کننده داشته‌است و در آزمون‌های آیلتس، هر ساله بیش از یک میلیون و دویست هزار نفر در ۱۲۰ کشور جهان شرکت می‌کنند. بنابراین، طراحی و برگزاری آزمون‌های مهارت از ضروریات گسترش زبان و ادبیات فارسی در دنیا است (Abbasi, 2014). در حال حاضر، آزمون‌های سامفا و آمفا به عنوان آزمون‌های مهارت زبان فارسی در ایران برگزار می‌شوند. متولی برگزاری آزمون سامفا، سازمان سنجش است که هر ساله این آزمون را در سه نوبت در ماه‌های خرداد، شهریور، و دی برگزار می‌کند. آزمون آمفا، آزمون مهارت بنیاد سعدی است که برای نخستین بار در سال ۱۳۹۶ برگزار شد.

سامفا نام اختصاری سنجش استانداردهای مهارت‌های فارسی، مهم‌ترین آزمون بسندگی در زبان فارسی است. این آزمون به سنجش استاندارد مهارت‌های زبانی افراد، بدون توجه به اینکه در کجا و با چه کتابی، زبان فارسی را فرا گرفته‌اند، می‌پردازد. سامفا در زبان فارسی، مانند آزمون‌های تافل و آیلتس در زبان انگلیسی است. این آزمون به کوشش مرکز همکاری‌های علمی بین‌المللی، سازمان سنجش آموزش کشور و اداره کل دانشجویان غیر ایرانی وزارت علوم، تحقیقات و فناوری و با همکاری استادان برجسته دانشگاه‌ها و موسسه‌های علمی کشور، طراحی و اجرا شده‌است. آزمون سامفا یکی از ملاک‌های تحصیل در دانشگاه‌های ایران است.

آزمون دیگر، آزمون مهارت زبان فارسی یا آمفا است. این آزمون ویژه فارسی‌آموزان در سراسر جهان و مورد تأیید شورای عالی انقلاب فرهنگی است. آزمون آمفا، دارای دو بخش عمومی و دانشگاهی است که بخش عمومی آن را بنیاد سعدی و بخش دانشگاهی آن را وزارت علوم، تحقیقات و فناوری ایران تهیه و تنظیم می‌کند.

آزمون اسپا (آزمون سنجش پارسی)، به عنوان نخستین آزمون مهارت زبان فارسی دانشگاه الزهراء نیز که به عنوان موضوع این پژوهش در نظر گرفته شده‌است، سومین آزمون طراحی شده است. آزمون بسندگی اسپا، نخستین آزمون بسندگی مرکز آرفای دانشگاه الزهراء است که در سال ۱۳۹۸ به سفارش موسسه خلاقان جوان (گروه آموزشی پارسیفن) طراحی و در میان فارسی‌آموزان این مرکز برگزار شد. به منظور بررسی روند طراحی، اجرا و کیفیت آزمون تدوین شده، مطالعه حاضر به سنجش پایایی و روایی محتوایی این آزمون مهارت زبان فارسی می‌پردازد.

۲. پیشینه پژوهش

به طور کلی آزمون‌سازی و به ویژه آزمون‌سازی زبان فرایند دشواری است. از آن جایی که آزمون‌ها برای گرفتن تصمیم‌هایی به کار می‌روند که بر زندگی مردم تأثیر می‌گذارند، باید تا اندازه‌ای ممکن اطلاعات دقیقی ارائه دهند تا آزمونگران بر مبنای آن‌ها تصمیمات عادلانه‌ای بگیرند. این امر، آزمون‌سازی را کاری بس دقیق می‌سازد (Farhadi, Jafarpur, & Birjandi, 2007). به همین سبب، در طراحی آزمون‌ها، به ویژه آزمون‌های بسندگی و مهارت، بررسی کیفیت و سوگیری آن‌ها از منظر پایایی^۱ و روایی^۲ اهمیت بسیاری دارد. بر مبنای دیدگاه بکمن^۳ (Bachman, 1995, p. 98)، نظریه‌های مطرح‌شده در حوزه پایایی، که به بررسی درستی نمره آزمون می‌پردازند، به سه دسته نظریه نمره واقعی کلاسیک^۴، نظریه تعمیم‌پذیری^۵ و نظریه پرسش و پاسخ^۶ گروه‌بندی می‌شوند. نظریه نمره واقعی کلاسیک، شامل مجموعه‌ای از فرض‌ها درباره ارتباط میان نمره واقعی یا ظاهری و عوامل تأثیرگذار بر این نمره‌ها است. نظریه تعمیم‌پذیری، رویکردی است که برای تخمین آنکه نمره یک فرد به عنوان نمونه‌ای برگرفته از جهانی از سنجش‌ها دارای چه مقدار تعمیم‌پذیری است. در نهایت، در نظریه سوم پرسش و پاسخ، آزمون از نظر سطح دشواری و در ارتباط با سطح توانایی فرد بررسی می‌شود. معروف‌ترین روش سنجش پایایی در آزمون‌ها، روش آلفای کرونباخ^۷ است که یکی از روش‌های سنجش پایایی بر مبنای نظریه نمره واقعی کلاسیک است.

روایی، مولفه دیگری است که در بررسی کیفیت آزمون‌ها مورد توجه قرار می‌گیرد. در بررسی روایی، فراتر از پایایی نمره‌های آزمون می‌رویم و روابط میان عملکرد آزمون و دیگر انواع عملکرد در بافت‌های دیگر را در نظر می‌گیریم. در بررسی روایی به صحت تعبیر و استفاده آزمون با در نظر گرفتن عملکرد و بافت آزمون پرداخته می‌شود، و اینکه آیا آزمون آنچه را که باید بسنجد، مورد ارزیابی قرار می‌دهد. به طور کلی، روایی دارای انواع گوناگونی اعم از روایی صوری^۸ (متمرکز بر شکل و صورت ظاهری آزمون)، محتوایی^۹ (با در نظر گرفتن پوشش و ربط محتوا)، ملاک-محور^{۱۰} (همزمان، پیش‌بین)، پیامدی^{۱۱} (تعیین پیامدهای مثبت یا منفی، خواسته یا

¹ reliability

² validity

³ Bachman

⁴ the classical true score theory (GTS)

⁵ generalizability theory (G-Theory)

⁶ item-response theory (IRT)

⁷ Cronbach's alpha

⁸ face validity

⁹ content validity

¹⁰ criterion-related validity

¹¹ consequential validity

ناخواسته)، و سازه‌ای^۱ (سنجش روابط بین سازه‌ها و نمره‌های آزمون) است. بسته به نوع و هدف آزمون، از روش‌های متفاوتی برای تایید روایی آزمون بهره گرفته می‌شود (Bachman, 1995, p. 82).

پژوهش‌های گوناگونی در پیوند با بررسی سنجش پایایی و روایی آزمون‌های متفاوت با روش‌های گوناگون، از قبیل آزمون‌های کلاسی یا استاندارد، انجام شده‌است. ایتو (Ito, 1996) در مطالعه خود به دنبال بررسی روایی و پایایی سوال‌های آزمون بسندگی انگلیسی در آزمون‌های ورودی دانشگاه‌های ژاپن بود. این بررسی در میان ۱۰۰ دانشجوی سال اولی انجام شد. این آزمون به سنجش مهارت‌های خواندن، نوشتن، شنیدن، دستور، و گفتار می‌پرداخت. یافته‌های این بررسی نشان داد که این آزمون، روایی و پایایی لازم را دارد. اگرچه پرسش‌های دستور تا اندازه‌ای از پایایی و روایی کافی برخوردار نبودند. آزمون تلفظ نیز از پایایی لازم برای سنجش مهارت شنیدن آزمودنی‌ها برخوردار نبود.

اوزر و همکاران (Ozer et al., 2014) در پژوهشی دیگر به بررسی پایایی و روایی محتوایی آزمون بسندگی زبان انگلیسی در میان زبان‌آموزان ترک پرداخته‌اند که انگلیسی را به عنوان زبان دوم می‌آموختند. یافته‌های بررسی آماری آن‌ها نشان داد که برخی از پرسش‌های آزمون نیازمند بازنگری هستند. رحمت و همکاران (Rahmat et al., 2015) نیز در پژوهش خود به بررسی دستاوردهای زبانی دانش‌آموختگان پنج دانشکده در مالزی پرداختند. هدف آن‌ها بررسی ارتباط بین آزمون بسندگی ورودی این دانشگاه با دستاوردهای دانش‌آموختگان آن است. این بررسی از طریق ارائه پرسشنامه و تحلیل آماری مربوط به آن در میان دانش‌آموختگان انجام شد. یافته‌های این پژوهش نشانگر آن بودند که حداقل نمره مورد پذیرش برای آزمون بسندگی باید کمتر از میزان در نظر گرفته شده باشد.

کونان و کارا (Kunnan & Carr, 2017) نیز در پژوهشی به مطالعه مقایسه‌ای میان آزمون بسندگی عمومی زبان انگلیسی و آزمون‌های پیشرفته و آزمون تافل اینترنتی پرداختند. آن‌ها در این مطالعه، به بررسی مقایسه‌ای مهارت‌های خواندن و نوشتن در این دو آزمون پرداختند. یافته‌های این بررسی نشان داد که اگر چه این دو آزمون ساختار خواندن یکسانی را ارزیابی می‌کنند، اما بر جنبه‌های مختلف آن تأکید دارند. یویون و میلینگ (Yuyun & Meyling, 2018) در پژوهش خود به بررسی آزمون بسندگی انگلیسی در مهارت‌های شنیدن، دستور، خواندن، و نوشتن در میان دانشجویان سال اول پرداختند. یافته‌های آن‌ها نشان داد که خواندن، مهارتی مشکل‌ساز برای این زبان‌آموزان است. قوامی لاهیج و همکاران (Ghavami Lahij et al., 2020) نیز در پژوهش

¹ construct validity

خود در راستای سنجش درک خوانداری بزرگسالان به بررسی روایی و پایایی آزمون تعیین سطح خود پرداختند. یافته‌های آنها نیز روایی محتوایی آزمون و پایایی آن را تأیید کرد. آنها ابزار استفاده‌شده در پژوهش خود را به عنوان ابزاری مناسب برای سنجش سطح درک خوانداری بزرگسالان تشخیص داده و معتقدند که می‌توان از این ابزار در برنامه‌ریزی‌های آموزشی برای انتخاب محتوای آموزشی بهره گرفت.

اصطلاح تأثیر آزمون^۱، به تأثیری که آزمون‌ها بر یادگیری و آموزش دارند، گفته می‌شود. به باور هانگ و هوانگ (Hung & Huang, 2019) تا به امروز بسیاری از پژوهش‌ها نشان داده‌اند که آزمون‌ها بر محتوای تدریس، طراحی دوره، و شیوه‌های آموزشی کلاس تأثیر می‌گذارند. آنها در مطالعه خود به دنبال واکاوی تأثیرات آزمون مهارت بر یادگیری دانشجویان در یک برنامه درسی انگلیسی در دانشگاه بودند. آنها با بررسی روابط بین تأثیر آزمون و ویژگی‌های یادگیرنده‌ها مانند رشته، جنسیت، و سطح مهارت، این شکاف پژوهشی را از بین بردند. در این پژوهش، روی هم‌رفته ۶۹۴ دانشجو از رشته‌های مهندسی، بازرگانی، و زبان‌های خارجی در یک دانشگاه ملی در تایوان مورد بررسی قرار گرفتند. یافته‌ها نشان‌دهنده تأثیر برجسته تأثیر آزمون بر جنبه‌هایی مانند تصویر شخصی، انگیزه یادگیری، احساسات، و فرصت‌های شغلی آینده بود. افزون بر این، بین تأثیر آزمون و سطح مهارت از نظر آماری همبستگی قابل توجهی وجود داشت و دانشجویان دختر و پسر از نظر آماری هیچ تفاوتی با هم نداشتند، و اختلاف معناداری هم بین رشته‌های مختلف وجود نداشت.

در پیوند با پژوهش‌هایی که در ایران انجام شده است می‌توان به پژوهش عباسی (Abbasi, 2014) اشاره کرد که به ارائه طرحی برای تهیه و تدوین یک آزمون بسندگی فارسی می‌پردازد. او در بررسی خود پس از تعریف مفاهیم مرتبط با این حوزه، به انواع آزمون‌های زبان پرداخته و به صورت نظری طرحی را برای ساختن این آزمون‌ها ارائه می‌کند. قنسولی (Ghonsoli, 2010) نیز در پژوهش خود پیرامون طراحی و روایی سنجی آزمون بسندگی فارسی، پس از ارائه تعریفی از مفاهیم این حوزه به بیان اهمیت این آزمون پرداخته و مراحل طراحی یک آزمون بسندگی را از طراحی و تدوین چارچوب کلی پرسش‌های آزمون تا بررسی‌های آماری آن بیان کرد. قنسولی (همان) این بررسی را در مقیاسی بسیار کوچک و به عنوان یک بررسی موردی انجام داده است. لازم به گفتن است که آزمودنی‌های این پژوهش سه دانشجوی خارجی دوره دکتری زبان و ادبیات فارسی هستند. پایایی آزمون در این پژوهش، از طریق آلفای کرونباخ

¹ washback

سنجیده شده است. در نتیجه گیری این پژوهش آمده است که این اثر را باید به عنوان اولین گام های طراحی و روایی سنجی یک آزمون بسندگی ملی در نظر گرفت که در آینده با اصلاحات و افزایش تعداد شرکت کنندگان می توان گام های موثرتری را در طرح چنین آزمونی برداشت. در این پژوهش نیز به مراحل طراحی و اجرای یک آزمون بسندگی فارسی در مقیاس وسیع تر پرداخته شده است.

در پژوهشی دیگر، زین العابدین و جمیل (Zainal Abidin & Jamil, 2015) به توصیف فرایند طراحی یک آزمون بسندگی جدید زبان انگلیسی برای دانشجویان خارجی که قصد ادامه تحصیل در مالزی را داشتند، پرداختند. نتایج بررسی آن ها نشان داد برای اینکه آزمونی روایی داشته باشد، باید ملاحظاتمانند تعداد آزمودنی ها، شرایط آزمون، و حوزه ای که آزمون در آن به کار می رود، در نظر گرفته شود. این آزمون در حقیقت یک آزمون بسندگی جدید با نام گیت^۱ بود. این پژوهش نشان داد که در طراحی آزمون های این چنینی جنبه های طراحی آزمون، اجرا، تحلیل، و روایی آن از اهمیت بسیاری برخوردارند. صحرائی و رضاپور (Sahraee & Rezapour, 2020) در پژوهش خود به بررسی روایی و پایایی آزمون تعیین سطح زبان فارسی بنیاد سعدی پرداختند. این آزمون از سه بخش اصلی دانش زبانی، نگارش، و صحبت کردن تشکیل شده است که بخش دانش زبانی شامل سه زیربخش دستور، واژه، و خواندن است. در این پژوهش، برای محاسبه پایایی از ضریب آلفای کرونباخ و برای بررسی روایی، از روایی سازی بهره گرفته شد. نتیجه این پژوهش نشان داد که آزمون در بخش ها و زیربخش های گوناگون و کلیت آزمون، پایایی مناسب و قابل قبولی را داشت، ولی با مقایسه میزان همبستگی بین دانش زبانی و مهارت های صحبت کردن و نوشتن می توان چنین نتیجه گرفت که زبان آموزان در بخش مهارت های تولیدی ضعیف عمل کرده اند و این نشان می دهد که در تدریس آن ها فقط به دانش محض زبانی پرداخته شده است و به دانش روندی توجه نشده است.

برای گسترش پژوهش ها و دستیابی به یافته های پژوهشی بیشتر در این حوزه، این مقاله به بررسی مراحل طراحی و اجرای یک آزمون بسندگی زبان فارسی و سنجش پایایی و روایی محتوایی آن پرداخته است.

۳. روش شناسی

پژوهش حاضر به بررسی روند طراحی یک آزمون بسندگی فارسی در مرکز آموزش زبان فارسی

^۱ GATE

دانشگاه الزهرا و اجرای آن در یک موسسه آموزشی می‌پردازد. فارسی‌آموزان شرکت‌کننده در این آزمون ۴۱ نفر از دانشجویان دوره‌های آموزش زبان فارسی موسسه خلاقان جوان (گروه آموزشی پارسیفن) در شهر مشهد هستند. این آزمون به سیله یک گروه از کارشناسان آموزش زبان فارسی به غیر فارسی‌زبانان در مرکز آرفای دانشگاه الزهرا طراحی و اجرا شده‌است. پایایی این آزمون به روش آماری سنجیده شده و برای سنجش روایی آن هم از دیدگاه‌های ۲۰ متخصص آموزش زبان فارسی استفاده شده‌است.

۴. مراحل طراحی و اجرای آزمون

برای طراحی این آزمون از مدل چندبعدی زبان که به مدل‌های استاندارد کنونی مانند تافل یا آیلتس شباهت دارد، استفاده می‌شود. پس از تشکیل یک کارگروه خبره ۶ نفره از اعضا برای طراحی آزمون و مشخص کردن افراد مناسب برای طراحی پرسش‌های هر مهارت، ابتدا به بررسی استانداردهای رعایت‌شده در طراحی آزمون‌های تافل و آیلتس به عنوان مدل‌های استاندارد طراحی سوال پرداخته شد. در ادامه، پس از صحبت‌های کارشناسی جدول مربوط به تعداد پرسش‌ها، نمره، و زمان هر بخش به صورت جدول (۱) طراحی شد.

جدول ۱: ساختار آزمون سنجش زبان پارسی

نام بخش	تعداد پرسش	نمره	زمان به دقیقه
۱ شنیدن	۱۸	۱۸	۲۰
۲ دستور و واژگان	۲۴	۲۴	۲۵
۳ خواندن و درک مطلب	۱۸	۱۸	۳۰
۴ نوشتن	۲	۲۰ (نگارش ۱: ۸ نمره) (نگارش ۲: ۱۲ نمره)	۵۵
مجموع	۶۲	۸۰	۱۳۰

سپس افراد خبره در گروه‌های تک‌نفره یا دونفره مسئول طراحی پرسش‌ها شدند. پس از طراحی پیش‌نویس اولیه، پرسش‌ها در جلسه‌های مختلف مورد بحث و تبادل نظر اعضا قرار گرفته و اصلاح و ویرایش شدند. این پرسش‌ها به ترتیب در بخش‌های شنیدن، دستور و واژگان، خواندن و درک مطلب، و نوشتن در یک دفترچه چیده شده که در مجموع شامل ۶۲ پرسش بودند. این

دفترچه در روز آزمون، در اختیار آزمون‌دهندگان قرار گرفت. چیدمان این پرسش‌ها از آسان به سخت و از سطح مقدماتی به پیشرفته در هر مهارت بوده است. به این ترتیب که یک سوم پرسش‌های ابتدایی هر بخش مربوط به سطح مقدماتی، یک سوم پرسش‌های میانه، مربوط به سطح متوسط، و یک سوم پایانی، پرسش‌های سطح پیشرفته بوده‌اند.

۸۰ نمره از ۱۰۰ نمره آزمون به پرسش‌های بالا اختصاص یافت و ۲۰ نمره باقی‌مانده هم برای آزمون شفاهی در نظر گرفته شد. شیوه‌نامه آزمون شفاهی در دفترچه دوم در اختیار مصاحبه‌کنندگان قرار گرفت. در پایان، نمره‌های فارسی‌آموزان در جدول زیر گزارش شد.

جدول ۲: گزارش نمرات در آزمون سنجش زبان پارسی به تفکیک مهارت‌ها و مؤلفه‌ها

بخش	نمره کل
شنیدن	۱۸
دستور و واژگان	۲۴
خواندن و درک مطلب	۱۸
نوشتن	۲۰
صحبت کردن	۲۰
مجموع	۱۰۰

۴. ۱. پرسش‌های شنیدن

در این آزمون، بخش شنیدن شامل سه بخش است که در بخش نخست آن از زبان آموزان خواسته می‌شود که پس از شنیدن فایل صوتی، اشیاء موجود در متن پرسش را با توجه به فایل صوتی پخش شده شماره‌گذاری کنند. بخش دوم شامل ۶ پرسش چهارگزینه‌ای است که با توجه به گفت‌وگوی پخش شده دوم، پاسخ داده می‌شود. در بخش سوم ۶ پرسش چهارگزینه‌ای دیگر تعریف شده است که با توجه به گفت‌وگوی سوم پاسخ داده می‌شود. لازم به گفتن است که پرسش‌های این بخش از آسان به سخت و از سطح مقدماتی تا پیشرفته در سه بخش تنظیم و نوشته شده‌اند. این پرسش‌ها از شماره یک تا هجده چیده شده‌اند.

۴. ۲. پرسش‌های واژگان و دستور

پرسش‌های این بخش، از شماره ۱۹ تا ۴۲ سازماندهی شده است و در قالب ۹ سوال چهارگزینه‌ای و سه متن کوتاه آزمون تکمیل متن^۱ طراحی شده است که از آسان به سخت مرتب شده‌اند. در این

^۱ cloze test

بخش، در متن‌هایی که از پرسش ۲۸ تا ۴۲ آمده‌اند، جای واژگان مورد سؤال خالی گذاشته شده و سپس از آزمودنی‌ها خواسته شده که از بین ۴ گزینه ارائه‌شده برای هر جای خالی واژه مناسب را انتخاب کنند.

۴.۳. پرسش‌های خواندن و درک مطلب

این بخش شامل ۳ متن خواندن است که از آسان به سخت و از سطح مقدماتی تا پیشرفته مرتب شده‌اند. بخش نخست، شامل قسمت‌های «الف تا ه» و یک متن کوتاه خواندن است که به ۵ بخش یک خطی دسته‌بندی شده‌است و سه پرسش چهارگزینه‌ای و یک سوال مربوط به دسته‌بندی‌های پاراگراف دارد (الف تا ه). در این بخش از آزمودنی‌ها خواسته شده مشخص کنند هر عبارت به کدام پاراگراف مربوط است و آن را از شماره ۴۱ تا ۴۶ در پاسخ نامه بنویسند.

بخش دوم شامل ۷ آگهی کوچک است که درباره نقاط گردشگری کشور نوشته شده‌اند. پرسش‌های مربوط به این بریده‌های آگهی از ۴۹ تا ۵۴ نوشته شده‌اند. عنوان این پرسش از این قرار است: «به هفت آگهی بالا مربوط به سفر در ایران نگاه کنید. کدام یک از تورها مناسب‌ترین گزینه برای افراد زیر است؟». در ادامه در صورت سوال‌ها ۶ موقعیت مختلف را بیان داشته و آزمودنی‌ها با شرایط مطرح شده در صورت سوال‌ها باید تور مناسب را برای هر موقعیت مطرح شده انتخاب کنند.

بخش سوم، متنی تا اندازه‌ای طولانی است. در این بخش، ۵ پاراگراف درباره یوزپلنگ ایرانی آمده‌است که در صورت سؤال آن از زبان آموز خواسته شده در هر جای خالی، یک تا سه واژه از متن را پیدا کنند و پاسخ‌های خود را در پاسخ‌نامه از شماره‌های ۵۵ تا ۶۶ بنویسند.

۴.۴. نوشتن

بخش نگارش شامل دو بخش است که بخش اول شامل یک تصویر است که در آن از آزمودنی خواسته می‌شود که تصویر را در قالب متنی شامل ۱۳۰ تا ۱۵۰ واژه توصیف کند. مدت زمان این بخش ۲۰ دقیقه است. در بخش دوم نگارش، با الگوگیری از آزمون آیلتس، به شرکت‌کنندگان یک موضوع داده می‌شود تا متنی شامل ۲۵۰ واژه درباره آن بنویسند. در این بخش آزمون، از شرکت‌کنندگان خواسته شده‌است که متنی درباره درس‌های آموزشی در مدرسه‌ها و به طور ویژه اهمیت درس تاریخ در مقایسه با درس‌های مربوط به علم و فناوری بنویسند و دلیل‌های خود را با آوردن تجربه‌های شخصی تقویت نمایند.

۴. ۵. پرسش‌های آزمون صحبت کردن

پرسش‌های آزمون صحبت کردن در دفترچه دوم در اختیار آزمون‌گران قرار گرفت. این دفترچه دارای دو بخش سؤال است. در بخش نخست جدول (۶) آمده‌است که در هر یک، ۶ پرسش مطرح شده‌است. در بخش دوم ۴ سوال تصویری بیان شده‌است. جدول پرسش‌ها در ۶ نسخه متفاوت، طراحی شده‌است تا پرسش‌ها متنوع باشند و از هر زبان آموز یک نسخه سؤال، از ابتدا تا انتها به طور کامل پرسش گردد. درجه دشواری در هر نسخه رعایت شده‌است و پرسش‌ها از سطح مقدماتی شروع می‌شود و سپس به سطح میانی و پیشرفته می‌رسد. برای تعیین سطح پرسش‌ها از دیدگاه‌های کارشناسان آموزش زبان فارسی به غیر فارسی‌زبانان در مرکز آرفای دانشگاه الزهراء بهره گرفته شده‌است. پرسش‌ها پس از طراحی بر اساس استانداردهای طراحی سؤال، در جلسه‌هایی مورد ارزیابی و نقد قرار گرفته‌اند.

۵. یافته‌های آزمون

۵. ۱. برآورد میانگین‌ها

میانگین نمره‌های به‌دست‌آمده زبان‌آموزان در این آزمون در مهارت‌های مختلف به قرار زیر لازم به گفتن است که این نمره‌ها از ۱۰۰ نمره محاسبه شده‌است که ۶۰ نمره آن مربوط به بخش‌های شنیدن، واژگان و دستور، و خواندن بوده‌است. ۲۰ نمره مربوط به نوشتن و ۲۰ نمره آن هم مربوط به آزمون صحبت کردن بوده‌است. در جدول (۳)، اطلاعات تکمیلی در مورد تعداد پرسش‌ها، مجموع نمره‌های آزمودنی‌ها، و میانگین نمره‌های زبان‌آموزان خارجی در مهارت‌های مختلف آزمون سنجش زبان پارسی ارائه شده‌است. در پایان، نیز مجموع میانگین‌ها در قالب یک نمره آورده شده‌است.

جدول ۳: میانگین نمرات زبان‌آموزان خارجی در مهارت‌های مختلف آزمون سنجش

زبان پارسی

شنیدن	واژگان و دستور	خواندن	نوشتن	صحبت کردن	
۱۸	۲۴	۱۸	۲	۷	تعداد سوالات
۵۴۶	۶۰۸	۴۲۵	۲۲۸.۲۵	۵۰۳	مجموع نمرات آزمودنی‌ها
۱۳.۳۱	۱۴.۸۲	۱۰.۳۶	۵.۵۶	۱۲.۲۶	میانگین نمره آزمودنی‌ها
۵۶.۳۱					مجموع میانگین‌ها

۵.۲. سنجش پایایی آزمون آسیا

پایایی یک وسیله سنجش درستی آزمون است و عمدتاً به دقت یافته‌های به دست آمده از آن اشاره می‌کند. پایایی به دقت، اعتمادپذیری، ثبات، یا تکرارپذیری نتایج آزمون اشاره می‌کند. روش‌های سنجش پایایی بر مبنای نظریه نمره واقعی کلاسیک متنوع اند. از جمله این روش‌ها، روش بازآزمایی (آزمون-بازآزمون)، روش فرم‌های موازی، روش دو نیمه کردن، روش کودر-ریچاردسون، روش ضریب آلفای کرونباخ است. ضریب آلفای کرونباخ که در این پژوهش برای تعیین پایایی آزمون مورد استفاده قرار گرفته است، یک روش تعیین پایایی آزمون است که در آن اجزا یا بخش‌های آزمون، برای سنجش ضریب پایایی آزمون به کار می‌روند. به منظور سنجش پایایی در این آزمون، نمره‌های فارسی آموزان در تمام مهارت‌های مورد سنجش آزمون وارد محیط برنامه اس.پی.اس.اس^۱ شد و در نهایت خروجی زیر به دست آمد.

جدول ۴: گزارش نمره‌های آزمودنی‌ها در محیط برنامه اس.پی.اس.اس

	Listening	Vocab	Reading	Writing	Oral
1	14.00	21.00	11.00	10.00	1.00
2	13.00	21.00	9.00	4.50	12.50
3	13.00	12.00	10.00	3.00	12.00
4	12.00	19.00	12.00	6.00	15.00
5	16.00	17.00	10.00	7.00	13.00
6	16.00	11.00	11.00	1.00	11.50
7	16.00	20.00	15.00	10.50	16.50
8	15.00	18.00	12.00	2.50	12.50
9	15.00	9.00	10.00	10.00	17.50
10	8.00	11.00	12.00	1.00	13.50
11	14.00	16.00	10.00	8.00	17.50
12	7.00	8.00	8.00	.00	11.00
13	16.00	13.00	11.00	11.00	11.00
14	17.00	11.00	8.00	1.00	3.50
15	14.00	18.00	10.00	4.50	13.50
16	12.00	16.00	9.00	12.00	17.50
17	16.00	9.00	10.00	4.00	11.25
18	12.00	19.00	13.00	.50	8.50
19	14.00	20.00	13.00	5.50	12.50
20	13.00	12.00	14.00	10.00	13.50
21	11.00	16.00	10.00	8.00	17.00
22	12.00	11.00	12.00	6.50	14.50
23	16.00	19.00	14.00	8.00	15.50

¹ SPSS

	Listening	Vocab	Reading	Writing	Oral
24	14.00	13.00	8.00	5.50	15.50
25	16.00	13.00	9.00	10.00	17.50
26	16.00	20.00	12.00	6.50	17.00
27	18.00	18.00	12.00	8.00	14.00
28	9.00	10.00	8.00	.75	10.50
29	12.00	11.00	9.00	.00	15.50
30	16.00	19.00	11.00	7.00	12.00
31	16.00	18.00	12.00	11.00	16.00
32	13.00	18.00	13.00	4.00	12.00
33	18.00	21.00	13.00	10.50	14.50
34	5.00	11.00	2.00	3.00	8.00
35	11.00	19.00	11.00	1.00	7.50
36	14.00	20.00	14.00	8.00	11.00
37	13.00	9.00	9.00	.00	7.00
38	4.00	7.00	1.00	.00	8.50
39	9.00	12.00	9.00	5.50	10.75
40	13.00	7.00	9.00	.00	.00
41	17.00	15.00	9.00	13.00	14.00

خروجی یافته‌های این روش به شرح زیر است.

جدول ۵: خلاصه پردازش موردی

%	N		
۱۰۰	۴۱	معتبر	موارد
.۰	۰	پاسخ داده نشده*	
۱۰۰	۴۱	کل	

جدول (۵) تعداد و درصد موارد معتبر، موارد پاسخ داده نشده، و کل موارد را نشان می‌دهد. با توجه به این جدول، تعداد نمونه ۴۱ نفر بوده و پرسش‌های بدون پاسخ هم در مطالعه مورد بررسی، وجود نداشته‌است. جدول (۶)، گزارش آمار توصیفی مهارت‌ها را ارائه می‌کند.

جدول ۶: آمار توصیفی

نمره میانگین/تعداد	پرسش‌های آزمون/تعداد
۱۳.۳۱	شنیدن
۱۴.۸۲	واژگان/دستور
۱۰.۳۶	خواندن
۵.۵۶	نوشتن
۱۲.۲۶	صحبت کردن
۴۱	تعداد شرکت کنندگان

میانگین مهارت شنیدن ۱۳/۳۱، دستور و واژگان ۱۴/۸۲، خواندن ۱۰/۳۶، نوشتن ۵/۵۶، و شفاهی ۱۲/۲۶ در نمونه مورد بررسی بوده است.

جدول ۷: گزارش تحلیل پایایی آزمون

آلفای کرونباخ استاندارد شده ^۱ پایه		
تعداد پرسش‌ها	پرسش‌ها	آلفای کرونباخ
۵	.۷۸	.۷۷

جدول (۷) آلفای کرونباخ و تعداد پرسش‌ها را بررسی می‌کند. از آن جایی که در این مطالعه تعداد بخش‌های آزمون ۵ مورد است (شنیدن، واژگان و دستور، خواندن و درک مطلب، نوشتن، و آزمون صحبت کردن) و مقدار آلفای کرونباخ ۰/۷۷ بوده که بزرگتر از ۰/۷ و نزدیک به ۰/۸ است، پس آزمون از پایایی قابل قبول و خوبی برخوردار است. اکنون، اگر بخواهیم اثر حذف مورد معینی را بر ضریب آلفای کرونباخ بررسی کنیم، در این صورت جدول زیر را خواهیم داشت.

جدول ۸: تحلیل آماری پرسش-کل

Cronbach's Alpha If Item Deleted	Squared Multiple Correlation	Corrected Item-Total Correlation	Scale Variance If Item Deleted	Scale Mean If Item Deleted	پرسش‌های آزمون
.71	.46	.57	132.12	43.03	شنیدن
.72	.41	.55	114.88	41.51	واژگان و دستور
.71	.50	.61	138.97	45.98	خواندن
.69	.47	.63	115.97	50.78	نوشتن
.77	.28	.40	129.78	44.07	صحبت کردن

در ستون نخست جدول بالا از راست، پرسش‌های آزمون، در ستون دوم، میانگین مجموع پرسش‌ها در صورت حذف پرسش مورد نظر، در ستون سوم واریانس سؤال‌ها، در ستون چهارم همبستگی پرسش با همه پرسش‌های آزمون، و در ستون آخر ضریب آلفای کرونباخ در صورت حذف پرسش مورد نظر نشان داده شده است. برای نمونه، با حذف پرسش بخش پنجم یعنی

¹ Cronbach's Alpha Based Standardized

صحبت کردن، ضریب آلفای کرونباخ به ۰/۷۷ افزایش می‌یابد.

۵.۳. سنجش روایی آزمون

ابزارهایی که برای گردآوری داده‌ها به کار می‌روند، باید از روایی یا اعتبار^۱ برخوردار باشند. روایی به این معنا است که روش یا ابزار به کاررفته تا چه اندازه می‌تواند خصوصیت مورد نظر در آزمون را درست اندازه‌گیری کند. رابطه بین روایی و پایایی از این قرار است که یک آزمون باید پایا باشد تا بتواند روایی داشته باشد. اگر آزمونی در هر بار اجرا روی تعدادی نمونه، یافته‌های گوناگونی به دست بدهد، آن آزمون یک آزمون پایا نخواهد بود و در واقع هیچ چیز را به درستی اندازه نخواهد گرفت، و اگر یک آزمون چیزی را به درستی اندازه‌گیری نکند، هیچ اطلاعات مفیدی به ما نخواهد داد. برای نمونه، یک آزمون ریاضی برای اندازه‌گیری محتوا و هدف‌های درس تاریخ، روا (مناسب) نیست. ولی این آزمون می‌تواند مطالب ریاضی را که اندازه می‌گیرد با دقت (به طور پایا) اندازه‌گیری کند. پس برای اینکه یک آزمون روا باشد، باید نخست پایا باشد. یعنی پایایی پیش شرط روایی است، و روایی آزمون کامل‌کننده پایایی است.

۵.۳.۱. تحلیل روایی محتوایی آزمون آسیا

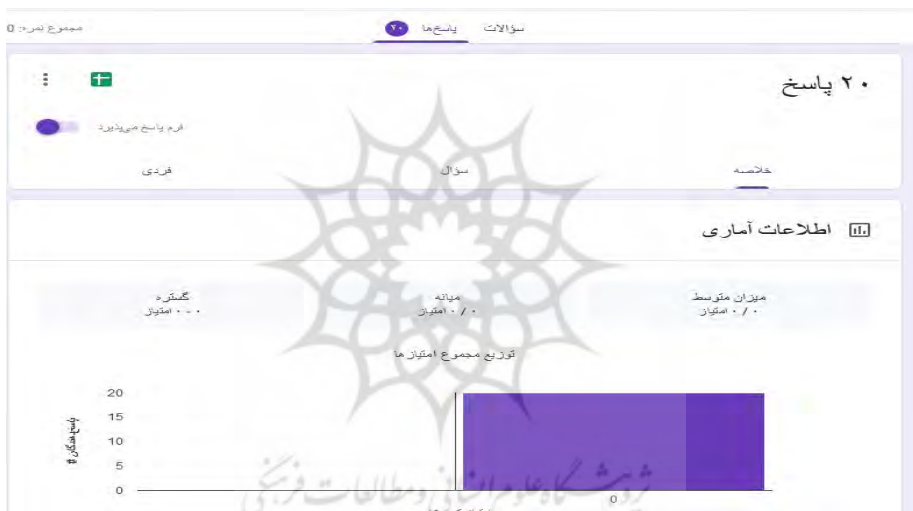
به دلیل محدودیت در تعداد شرکت‌کنندگان در آزمون که از اندازه معمول در سنجش‌های کمی کمتر بوده‌است، ناچار به استفاده از تحلیل روایی محتوایی در تعیین ارتباط پرسش‌های آزمون با سازه‌های دانش زبانی آموخته‌شده آزمودنی‌ها شدیم. یکی از روش‌های انجام تحلیل روایی محتوایی، استناد به دیدگاه‌های کارشناسان هر رشته است. در این راستا، ابتدا یک فرم نظرخواهی در گوگل فرم^۲ ساخته شد که در آن نمونه پرسش‌های هر مهارت در سطح‌های مختلف اعم از مقدماتی، میانی و پیشرفته به صورت تصویر در صفحات وب آورده شده و پس از طراحی فرم مورد نظر که در ادامه نمونه‌ای از آن آورده شده‌است (شکل ۱)، لینک این فرم نظرخواهی برای ۲۰ مدرس آموزش زبان فارسی به غیر فارسی‌زبانان فرستاده شده و از آن‌ها خواسته شد که نظر خود را در مورد مناسب یا نامناسب بودن پرسش‌ها از نظر زبان و محتوا (ارتباط، سادگی، وضوح) با آوردن دلیل بنویسند. از آن جایی که یک آزمون بسندگی باید در بر دارنده محتوای سه سطح مقدماتی، میانی و پیشرفته باشد، در طراحی آن به تعداد برابر برای هر سه سطح پرسش‌هایی در نظر گرفته شده بود تا سطح آزمودنی در پاسخ به پرسش‌ها تعیین شود. در این راستا و برای سنجش

¹ validity

² Google Form

روایی پرسش‌ها، نمونه‌هایی از آزمون در هر سه سطح در قالب فرم نظرسنجی در اختیار کارشناسان آموزش زبان فارسی که جمعی از مدرسان باسابقه از مراکز مختلف آزفا بوده‌اند، قرار داده شد. فرم نظرخواهی طراحی شده بر روی گوگل فرم دارای ۶ صفحه بود که هر صفحه به یک مولفه زبانی، اختصاص داده شده بود. در ادامه، فرم نظرسنجی ارائه می‌شود. پس از گردآوری فرم‌ها به بررسی نظرها مطابق نمونه‌های ارائه شده پرداخته شد. ۲۰ نفر از کارشناسان آموزش زبان فارسی به غیر فارسی‌زبانان در این نظرسنجی شرکت کرده‌اند.

شکل ۱: اطلاعات فرم نظرسنجی از گروه کارشناسان در مورد تحلیل محتوایی آزمون آسیا در گوگل فرم



این افراد پس از دریافت لینک باید در صفحه اول پس از وارد کردن مشخصات از قبیل نام و نام خانوادگی، میزان سابقه تدریس دوره‌های آموزش فارسی به غیر فارسی‌زبانان و ایمیل خود وارد فرایند پاسخ‌گویی و اعلام نظر خود نسبت به پرسش‌های آزمون شوند.



شکل ۲: نمونه فرم نظرخواهی از گروه کارشناسان بر روی گوگل فرم

نمونه سؤالهای هر مهارت به صورت یک تصویر از فایل پرسش در هر صفحه از گوگل فرم
قرار داده شد و در پایان از مدرسان خواسته ایم که نظرات خود را در مورد مناسب بودن یا نبودن
سوالات با ارائه دلایل توجیهی بیان کنند.

واژگان و دستور

گروهی فرست را انتخاب کنید و فر پاسخنامه از شماره ۴۴-۴۹ تکلیف دارید.

۴۹. میرزا آسود و باگلار ۴۹ ساله است و باگلار نیز ۴ ساله است. ...

۴۵. ...

۴۶. ...

۴۷. ...

۴۸. ...

۴۹. ...

۵۰. ...

پاسخ شما

۳. لطفاً نظرات خود را در مورد طراحی و کیفیت سوالات لیکن مشخص بنویسید. به نظرات شما این سوالات برای سطح پیشرفته مناسب هستند؟ چرا؟

دوره دانش که جهانی کنونی است بر دانش که گذرات پنج سال پیش ...

۴۴. ...

۴۵. ...

۴۶. ...

۴۷. ...

۴۸. ...

۴۹. ...

۵۰. ...

پاسخ شما

۴. لطفاً نظرات خود را در مورد طراحی و کیفیت سوالات بخش خواندن و درک مطلب بنویسید. به نظرات شما این سوالات برای سطح متوسط مناسب هستند؟ چرا؟

شکل ۳: پرسش‌های آزمون و فرم نظرخواهی در مورد زبان و محتوای پرسش‌ها بر روی گوگل فرم

با توجه به تعداد بسیار صفحه‌های نظرسنجی، فقط چند نظر از میان نظرات مربوط به مهارت شنیدن به عنوان نمونه در ادامه آمده‌است.

۱. به نظر من این سؤالات مناسب سطح مقدماتی هستند.
۲. بله، می‌تواند متناسب با دوره مقدماتی باشد.
۳. بله، موضوع گفت‌وگو و واژگان منتخب، کاربردی و متناسب با سطح و تصویر گویا و روشن است.
۴. برای پایان سطح مناسب است. محتوای سوال و میزان دانش خواسته شده مناسب است اما شیوه طرح و ساختار سؤال کمی شلوغ و پیچیده است.
۵. مناسب است ولی با توجه به سطح، دستورالعمل‌ها نیاز به توضیح بیشتر دارد.

شکل ۴: نظرات مدرسان آموزش زبان فارسی به خارجی‌ان در گوگل فرم

نظرات مدرسان آموزش زبان فارسی به غیر فارسی زبانان در تمام مهارت‌ها به طور خلاصه و با اشاره به درصد در جدول (۹) آمده‌است.

جدول ۹: نظرات کارشناسان آموزش زبان فارسی به غیر فارسی زبانان دربارهٔ زبان و محتوای پرسش‌های آزمون بر مبنای درصد

مهارت	مناسب با سطح	نیازمند اصلاح	عدم تناسب با سطح
شنیدن	٪۶۵	٪۱۵	٪۲۰
واژگان	٪۵۰	٪۳۵	٪۱۵
دستور	٪۹۰	٪۵	٪۵
خواندن و درک مطلب	٪۴۰	٪۴۵	٪۱۵
نوشتن	٪۹۰	٪۱۰	۰
گفتار	٪۸۵	٪۵	٪۱۰

پس از گردآوری نظرات به دسته‌بندی آن‌ها در سه دسته پرسش‌های مناسب با سطح، پرسش‌های نیازمند تغییر با توجه به سطح، و پرسش‌های نامناسب برای سطح تعریف شده دسته‌بندی شدند. همان گونه که این جدول نشان می‌دهد، برای نمونه ٪۶۵ از کارشناسان شرکت کننده در نظرسنجی سطح پرسش‌های شنیدن را مناسب ارزیابی کرده‌اند. ٪۱۵ آن را نیازمند تغییر، و ٪۲۰ هم آن را برای سطوح در نظر گرفته شده نامناسب دانسته‌اند. به تفکیک مهارت‌ها، بررسی روایی محتوایی این آزمون نشان داد که مهارت‌های دستور و نوشتن با کسب ٪۹۰ از نظر مساعد کارشناسان، بالاترین میزان و در پرسش‌های مربوط به مهارت خواندن و درک مطلب کمترین

روایی محتوایی را داشته‌است. به طور کلی، با در نظر گرفتن تعداد کارشناسان ارزیاب محتوای آزمون و تأیید و توافق بیشتر آن‌ها در مورد پوشش و ارتباط محتوا، شاخص نسبت روایی محتوایی آزمون (Lawshe, 1975) که یکی از رایج‌ترین روش‌های سنجش کمی روایی محتوایی است، نمایانگر روایی محتوایی بالای پرسش‌های آزمون در تمام مهارت‌ها است. بر مبنای جدول لاوشه، نسبت روایی محتوایی برای ۲۰ نفر کارشناس آموزش زبان فارسی به خارجی‌ان ۰/۴۲ است که در این مقاله برای ۶۲ پرسش در محدوده (۰/۵-۱) قرار گرفت.

۶. بحث

این پژوهش به بررسی طراحی آزمون بسندگی آسیا پرداخت و میزان پایایی و روایی محتوایی این آزمون را مورد سنجش قرار داد. در طراحی این آزمون پس از تشکیل یک کارگروه ۶ نفره از کارشناسان آموزش زبان فارسی به خارجی‌ها و مشخص کردن افراد مناسب برای طراحی پرسش‌های هر مهارت، ابتدا به بررسی استانداردهای رعایت‌شده در طراحی آزمون‌های تافل و آیلتس به عنوان مدل‌های استاندارد طراحی سوال آزمون استانده پرداخته شد. سپس، پس از طراحی پرسش‌های آزمون، این سوالات در جلسه‌هایی، مورد تحلیل و بررسی دیگر اعضا قرار گرفت. پس از انجام بررسی‌های لازم، آزمونی با ۱۸ سوال شنیدن، ۲۴ سوال دستور و واژگان، ۱۸ سوال خواندن و درک مطلب، و ۲ سوال نوشتن شد که جزئیات آن در جدول (۱) گزارش شده‌است، طراحی شد. این پرسش‌ها در مجموع ۸۰ نمره از آزمون را در بر می‌گرفتند. ۲۰ نمره باقی‌مانده مربوط به آزمون صحبت کردن می‌شد که در دفترچه جداگانه‌ای به صورت مصاحبه از زبان‌آموزان پرسیده شد. شرح پرسش‌های آزمون و چیدمان آن‌ها در مقاله آمده‌است.

پایایی این آزمون به وسیله آلفای کرونباخ سنجیده شد و نتیجه نمایانگر این بود که آزمون از پایایی قابل قبول و خوبی برخوردار است. برای سنجش روایی این آزمون از شیوه روایی محتوایی، با استناد به نظر کارشناسان آموزش زبان فارسی به غیر فارسی‌زبانان (تعداد، 20 نفر) استفاده شد. در این راستا، ابتدا یک فرم نظرسنجی در ۶ صفحه در گوگل فرم طراحی شد و لینک آن در اختیار مدرسان قرار گرفت. پس از گردآوری فرم‌های نظرسنجی به دسته‌بندی نظرات در سه قالب «مناسب با سطح»، «نیازمند تغییر»، و «نامناسب با سطح» پرداخته شد. نظرات کارشناسان آموزش زبان فارسی نیز نمایانگر برخورداری آزمون از روایی محتوایی مناسب بوده است. باتوجه به جمع‌بندی نظرات می‌توان نتیجه گرفت که تعداد بیشتری از کارشناسان ساخت و طراحی سوالات مهارت‌های آزمون را برای سطح‌های تعریف‌شده آن مناسب دانسته‌اند.

اگرچه گروهی که پرسش‌ها را نیازمند تغییر دانسته‌اند نیز با بیان دلایل خود بیان داشته‌اند که با تغییرات مورد نظر آن‌ها این پرسش‌ها مناسب‌سازی خواهند شد. تعداد کمی هم برخی پرسش‌ها را در برخی مهارت‌ها نامناسب دانسته‌اند. برای نمونه، در بخش پرسش‌های مهارت خواندن ۱۵٪ از شرکت‌کنندگان سطح پرسش‌ها را نامناسب ارزیابی کرده‌اند. بنابراین، به طور کلی با در نظر گرفتن مجموع نظرات می‌توان نتیجه گرفت که پرسش‌ها از نظر کارشناسان دارای روایی محتوایی بوده‌است. این روایی در مهارت‌های دستور و نوشتن با کسب ۹۰٪ از نظر مساعد کارشناسان، بالاترین میزان و در پرسش‌های مربوط به مهارت خواندن کمترین روایی را داشته‌است. میزان روایی این مجموعه پرسش‌ها از دید کارشناسان، از بهترین تا ضعیف‌ترین نیز به این ترتیب بوده‌است: دستور، نوشتن، گفتار (سوالات شفاهی)، شنیدن، واژگان، و خواندن و درک مطلب.

از بررسی‌های انجام شده می‌توان نتیجه گرفت که واژگان و دستور با کسب بالاترین میانگین نمره‌های آزمون دارای سطح آسانی بیشتر و نوشتن با کسب کمترین نمره میانگین از سطح سختی بالاتری برای زبان‌آموزان برخوردار بوده‌است. میزان سختی این مجموعه پرسش‌ها از آسان به سخت به این ترتیب بوده‌است: واژگان و دستور، شنیدن، پرسش‌های صحبت کردن، خواندن، نوشتن. این یافته‌ها می‌تواند نشان‌دهنده این باشد که آزمودنی‌ها نیازمند تمرکز بیشتری بر روی مهارت نوشتن هستند. این یافته‌ها خود می‌تواند تأییدکننده این مطلب باشد که آزمودنی‌ها در مهارت نوشتن دارای ضعف هستند و باید در شیوه آموزشی آن‌ها به این مهارت پایه‌ای اهمیت بیشتری داد. بنابراین پیشنهاد می‌شود که بر روی این مهارت با استناد به روش‌های جدید در حوزه رویکرد ارتباطی، از قبیل روش تکلیف-محور کار شود و تمرین‌های مربوط به تقویت این مهارت در آموزش آن‌ها گنجانده شود.

۷. نتیجه‌گیری

در این پژوهش، به بررسی چگونگی طراحی پرسش‌ها و اجرای آزمون آسپا به عنوان آزمون بسندگی فارسی، طراحی شده در دانشگاه الزهراء، پرداخته شده‌است. این آزمون به عنوان سومین آزمون مهارت‌سنجی زبان فارسی، در میان فارسی‌آموزان موسسه‌ای در شهر مشهد اجرا شده‌است. با توجه به سیاست محرمانه نگهداشتن پرسش‌های آزمون‌های سامفا و آمفا این پژوهش می‌تواند راهگشای طراحی آزمون‌های مشابه در مقیاس‌های کوچک و بزرگ باشد و به نوعی می‌توان آن را ادامه‌دهنده پژوهش‌هایی همچون اثر قنسولی (Ghonsoli, 2010) دانست که در آن نخستین گام‌های طراحی و اجرای این آزمون‌ها در مقیاس کلاسی و کوچک و با چند زبان‌آموز برداشته

شد.

در این مقاله که در سطح یک موسسه برای سنجش مهارت فارسی آموزان برگزار شد، آزمون طراحی شده در میان ۴۱ شرکت کننده اجرا شد. زبان آموزان برای کسب مدرک این آزمون به سنجش میزان مهارت خود در زبان فارسی پرداختند. برای بررسی پایایی این آزمون از آلفای کرونباخ استفاده شد. یافته‌های آماری نمایانگر پایایی خوب و مناسب این آزمون بود. روایی این آزمون هم به شیوه‌ی روایی محتوایی و با استفاده از نظر کارشناسان آموزش زبان فارسی به غیر فارسی‌زبانان سنجیده شد که نشان‌دهنده‌ی برخورداری روایی محتوایی بسیار خوبی است. این کارشناسان معتقد بودند که پرسش‌های دستور و نوشتن از بیشترین انطباق با سطح زبان آموزان برخوردار بوده‌اند و پرسش‌های خواندن و درک مطلب از کمترین میزان انطباق برخوردار بوده‌اند. به کوتاه سخن، می‌توان گفت که در این پژوهش، برخی از پرسش‌های آزمون نیازمند بازنگری هستند.

در پژوهشی دیگر، ایتو (Ito, 1996)، پس از بررسی روایی پرسش‌های آزمون خود به این نتیجه رسید که پرسش‌های دستور از روایی کافی برخوردار نیستند. هرچند که کل آزمون مانند نمونه پیش‌رو از روایی و پایایی کافی برخوردار بوده‌است. با در نظر گرفتن میانگین نمره‌های به دست آمده‌ی زبان آموزان هم می‌توان نتیجه گرفت که واژگان و دستور دارای سطح آسانی بیشتر و نوشتن از سطح سختی بالاتری برای زبان آموزان برخوردار بوده‌است. این در حالی است که در پژوهش یویون و میلینگ (Yuyun & Meyling, 2018)، خواندن مهارتی مشکل‌ساز برای آزمودنی‌ها بوده‌است.

می‌توان گفت که یافته‌های آزمون‌های بسندگی می‌تواند افزون بر تعیین سطح زبان آموزان، راهگشای برنامه‌ریزی‌های لازم برای آن‌ها نیز باشد. افزون بر این طراحی، اجرا و بررسی آزمون با استفاده از یافته‌های آن می‌تواند زمینه‌ساز ساخت آزمون‌های استاندارد در مراکز و مؤسسه‌های دیگر باشد. با وجود مراکز بسیاری که در چند سال اخیر اقدام به جذب و برگزاری کلاس‌هایی برای فارسی‌آموزانی از سراسر جهان داشته‌اند، وجود آزمون‌های بسندگی استاندارد در سطح‌های مختلف، از یک مؤسسه تا سطح‌های ملی می‌تواند جهت مناسبی به امر آموزش فارسی‌آموزان بدهد و تعیین سطح را از حالت تجربی و سلیقه‌محور به صورت آزمونی علمی و استاندارد درآورد. وجود آزمون‌های سامفا و آمفا در مقیاس‌های بزرگ به عنوان نقطه عطفی در این روند در نظر گرفته می‌شود، اما با توجه به بزرگی مقیاس این آزمون‌ها و امنیت بالای آن‌ها عملاً قابلیت استفاده در مقیاس‌های کوچک‌تر مانند یک مؤسسه یا مرکز را ندارند. بنابراین، طراحی آزمون‌های

بسندهای مانند آزمون اسپا می تواند نیازهای مراکز مختلف آموزش زبان فارسی را در این راستا برآورده سازد.

بر پایه تجربه به دست آمده از این پژوهش، طراحی یک آزمون بسندگی در تیمی از کارشناسان از مدرسان با تجربه آموزش زبان فارسی می تواند از طراحی مشخصات آزمون شروع شده و به اجرا و تعیین پایایی و روایی آزمون ختم شود. این آزمون ها می توانند در مقیاس های کوچک یا بزرگ اجرا شوند و در پژوهش های مانند مقاله حاضر مورد ارزیابی قرار گیرند.

با توجه به محدودیت دسترسی به شرکت کنندگان بیشتر، می توان این پژوهش را در باب آزمون بسندگی فارسی در مقیاس های بزرگ تر انجام داد تا امکان انجام تحلیل روایی سازه ای آن نیز با استفاده از روش های تحلیل عاملی اکتشافی و تأییدی (تحلیل مسیری) فراهم شود. امید است که طراحی و اجرای آزمون های بسندگی داخلی بتواند سبب برداشتن گام های موثری در آینده نزدیک برای برگزاری بین المللی آزمون های بسندگی فارسی در مقیاس های بسیار بزرگ تر شود.

فهرست منابع

براون، اچ، داگلاس (۱۹۸۰). اصول یادگیری و تدریس زبان. ترجمه مجدالدین کیوانی. تهران: مرکز نشر دانشگاهی.

صحرايي، رضامراد و غلامحسين رضاپور (۱۳۹۸). «ارزیابی روایی و پایایی آزمون تعیین سطح زبان فارسی برای غیرفارسی زبانان بنیاد سعدی». *اننازگیری تربیتی*. دوره ۱۰. شماره ۳۷. صص ۹۱-۱۱۱.

عباسی، زهرا (۱۳۹۳). «تهیه و تدوین آزمون بسندگی زبان فارسی». *مجموعه مقالات نخستین همایش آموزش زبان فارسی*. به کوشش انجمن علمی استادان زبان و ادبیات فارسی. تهران: دانشگاه تربیت مدرس. صص ۴۲۵-۴۳۳.

قنسولی، بهزاد (۱۳۸۹). «طراحی و روایی سنجی آزمون بسندگی زبان فارسی». *پژوهش زبان های خارجی*. دوره ۱۲. شماره ۵۷. صص ۱۱۵-۱۲۹.

قوامی لاهیج، سارانا. مریم دانای طوس، عبدالرضا تحریری و علی ربیع (۱۳۹۸). «طراحی و اجرای مقدماتی آزمون تعیین سطح درک خوانداری بزرگسالان: شواهدی از دانشجویان کارشناسی ارشد دانشگاه گیلان». *زبان پژوهی*. دوره ۱۱. شماره ۳۳. صص ۲۶۵-۲۹۷.

وکلیلی فرد، امیررضا (۱۳۹۴). «سخن سردبیر». *پژوهش نامه آموزش زبان فارسی به غیرفارسی زبانان*. دوره ۴. شماره ۱. ص ۱.

References

Abbasi, Z. (2014, March). Designing a Persian language proficiency test. In Scientific Association of Professors of Persian Language and Literature

- (Eds.), *The First Conference on Persian Language and Literature Teaching* (pp. 425-433). Tehran: Tarbiat Modares University [In Persian].
- Bachman. L.F. (1995). *Fundamental considerations in language testing*. Oxford: Oxford University Press
- Brown. H. D. (1980). *Principles of language learning and teaching*. Englewood Cliffs, N.J.: Prentice Hall.
- Farhadi, H., Jafarpour, A., & Birjandi, P. (2007). *Language skills testing: From theory to practice*. Tehran: SAMT [In Persian].
- Ghonsoli, B. (2010). Designing and validation of a Persian proficiency test. *The Journal of Research in Foreign Languages*, 57, 115-129. [In Persian].
- Ghavami Lahij, S. Danaye Tous, M. Tahriri, A., & Rabi, A. (2020). Developing and preliminary implementation of adults reading comprehension placement test: Evidence from Guilan university M.A. students. *Journal of Language Research*, 11(33), 265- 297. [In Persian].
- Hung, S. A., & Huang, H. T. (2019). Standardized proficiency tests in a campus-wide English curriculum: A washback study. *Language Testing in Asia*, 9, 1-17.
- Ito, A. (1996). Testing English tests: A language proficiency perspective. *JALT*, 18(2), 183-197.
- Kunnan, A., & Carr, N. (2017). A comparability study between the General English Proficiency Test: Advanced and the Internet-Based Test of English as a Foreign Language. *Language Testing in Asia*, 7(1), 1-16.
doi:10.1186/s40468-017-0048-x
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575. doi:10.1111/j.1744-6570.1975.tb01393.x
- Ozer, I., Fitzgerald, S., Sulbaran, E., & Garvey, D. (2014). Reliability and content validity of an English as a Foreign Language (EFL) grade-level test for Turkish primary grade Students. *Procedia-Social and Behavioral Sciences*, 112, 924-929.
doi:10.1016/j.sbspro.2014.01.1250
- Rahmat, N., Min, L., Sing, N., & Mior Yusup, F. (2015). English language proficiency tests and academic achievement: A study on the Malaysian University English Test as a predictor of technical program undergraduates' academic achievement. *Advances in Language and Literary Studies*, 6 (1), 114-119. doi:10.7575/aiac.all
- Sahraee, R.M, & Rezapour, Gh. (2020). Evaluation of Reliability and Validity of Persian for non-Persian Speakers Placement Tests of Sa'adi Foundation. *Quarterly of Educational Measurement*, 10 (37), 91-111. [In Persian].
- Vakilifard, A. R. (2015). Message from the editor in chief. *Journal of Teaching Persian to Speakers of Other Languages*, 4(9), 1-2 [In Persian].
- Zainal Abidin, S., & Jamil, A. (2015). Toward an English proficiency test for postgraduates in Malaysia. *SAGE Open*, 5(3), 1-10.
doi: 10.1177/2158244015597725
- Yuyun, I., & Meyling, M. (2018). A study of English proficiency test among the first-year university students. *Journal of Language and Literature*, 18(1), 1-8.

